Proportional Size Distribution (PSD)

Alex J. Benecke

January 10, 2018

I will use a Chi-Square test to see if there is a difference between PSD between years (2013 - 2017).

Data Preparation

Load Data

```
lmbs <- read.csv("Data/Clean-Data/largemouth-bass_Wr_Stock.csv") %>% filter(Year <
    2017) %>% arrange(Year, FID, Length)
lmbs$fyr <- factor(lmbs$fyr)
```

```
unique(lmbs$Year) ### See that there is no 2012
```

```
## [1] 2013 2014 2015 2016
```

View Data

```
(lmbs.LF <- xtabs(~Year+gcat,data=lmbs))
```

```
##       gcat
## Year   preferred quality stock
##   2013        16      40    41
##   2014        18      57    65
##   2015        15      38    14
##   2016        10      44    52
```

Chi-Squares Test

Is there a difference in the number of fish in each gabelhouse length category (PSD-X) during the years 2013 - 2017?

```
chisq.test(lmbs.LF)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmbs.LF
## X-squared = 17.48, df = 6, p-value = 0.007673
```

This seems to suggest that the proportional stock distribution (PSD) is different for largemouth bass between years ($X^2 = 17.48$, df = 6, P = 0.007673).

Where are the differences in PSD X-Y for each year?

Creating a table of percent of fish in each gabelhouse length interval (PSD X-Y) in each year.

```
round(prop.table(lmbs.LF,margin=1)*100,0)
```

```
##       gcat
## Year   preferred quality stock
##   2013        16      41    42
##   2014        13      41    46
##   2015        22      57    21
##   2016         9      42    49
```

Remarkably the percent of quality fish is the same for 2014 and 2016. the percent of quality length fish is very similar between 2013 (42%), 2014 (41%), 2016 (41%), and 2017 (43%). This trend is almost the same for stock length fish (2013 42%, 2014 46%, 2016 49%), however 2015 and 2017 have a smaller percentage of stock length fish (21% and 31% respectivley). This may be explained by more variability among years for this length category, reduced sampling efficiency for smaller fish, or unstable recruitment. The year 2015 seems to break the trend with a greater percentage of preferred (22%) and quality (57%) length individuals than other years.

1) Could this be some sort of sampling bias?
2) Could this be a result of sampling different sites? where the 2015 sites more suitable for LMB?

PSD-X Without 2013 - 2016 Count of Fish

Compare PSD-Q between years 2013 - 2016

```
lmbs %<>% mutate(gcatQ=mapvalues(gcat,
                                 from=c("stock","quality","preferred"),
                                 to=c("quality-", "quality+", "quality+")),
                 gcatQ=droplevels(gcatQ))

(lmb.LFQ <- xtabs(~Year+gcatQ,data = lmbs))
```

```
##       gcatQ
## Year   quality+ quality-
##   2013       56       41
##   2014       75       65
##   2015       53       14
##   2016       54       52
```

```
chisq.test(lmb.LFQ)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmb.LFQ
## X-squared = 15.552, df = 3, p-value = 0.001401
```

```
(ps.Q <- c(chisq.test(lmb.LFQ[c(1,2),])$p.value,  ### 2013-2014
           chisq.test(lmb.LFQ[c(1,3),])$p.value,  ### 2013-2015
           chisq.test(lmb.LFQ[c(1,4),])$p.value,  ### 2013-2016
           chisq.test(lmb.LFQ[c(2,3),])$p.value,  ### 2014-2015
           chisq.test(lmb.LFQ[c(2,4),])$p.value,  ### 2014-2016
           chisq.test(lmb.LFQ[c(3,4),])$p.value)) ### 2015-2016

(p.val.Q <- p.adjust(ps.Q))
```

```
## [1] "13-14" "13-15" "13-16" "14-15" "14-16" "15-16"
```

```
##    Year p-value Adjusted p
## 1 13-14  0.6167     1.0000
## 2 13-15  0.0073     0.0293
## 3 13-16  0.4073     1.0000
## 4 14-15  0.0007     0.0036
## 5 14-16  0.7796     1.0000
## 6 15-16  0.0004     0.0023
```

The PSD-Q of largemouth bass *is significantly different* for at least one of the years during 2013 - 2016 (Chi-Squared, $X^2 = 15.556$, df = 3, p = 0.0014). The adjusted p-values show a *significant difference* in PSD-Q between years **2013 - 2015** (p = 0.0293), **2014 - 2016** (p = 0.0036), and **2015 - 2016** (p = 0.0023). The PSD-Q *is not significantly different* between any other years.

Compare PSD-P between years 2013 - 2016

```
lmbs %<>% mutate(gcatP=mapvalues(gcat,
                                 from=c("stock","quality","preferred"),
                                 to=c("preferred-", "preferred-", "preferred+")),
                 gcatP=droplevels(gcatP))

(lmb.LFP <- xtabs(~Year+gcatP,data = lmbs))
```

```
##       gcatP
## Year   preferred+ preferred-
##   2013         16         81
##   2014         18        122
##   2015         15         52
##   2016         10         96
```

```
chisq.test(lmb.LFP)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmb.LFP
## X-squared = 6.2082, df = 3, p-value = 0.1019
```

```
(ps.P <- c(chisq.test(lmb.LFP[c(1,2),])$p.value,   ### 2013-2014
           chisq.test(lmb.LFP[c(1,3),])$p.value,   ### 2013-2015
           chisq.test(lmb.LFP[c(1,4),])$p.value,   ### 2013-2016
           chisq.test(lmb.LFP[c(2,3),])$p.value,   ### 2014-2015
           chisq.test(lmb.LFP[c(2,4),])$p.value,   ### 2014-2016
           chisq.test(lmb.LFP[c(3,4),])$p.value))  ### 2015-2016

(p.val.P <- p.adjust(ps.P))
```

```
## [1] "13-14" "13-15" "13-16" "14-15" "14-16" "15-16"
```

```
##    Year p-value Adjusted p
## 1 13-14  0.5504     1.0000
## 2 13-15  0.4565     1.0000
## 3 13-16  0.1958     0.7833
## 4 14-15  0.1212     0.6060
## 5 14-16  0.5258     1.0000
## 6 15-16  0.0325     0.1948
```

The PSD-P of largemouth bass is *not significantly different* for any years during 2013 - 2016 (Chi-Squared, $X^2 = 5.45$, df = 3, p = 0.14). The adjusted p-values show no difference in the PSD-P between years (2013 - 2016).

PSD-X 2013 - 2016 CPUE

```
cpe <- read.csv("Data/Clean-Data/CPUE-gcat_2013-2016.csv") %>%
  filterD(Species == 317) %>%
  filterD(gcat != "substock") %>%
  filterD(!is.na(gcat))


headtail(cpe)
```

```
##     Year Site    effort Species      gcat caught   cpe.hr
## 1   2013    2 0.2536111     317     stock      7 27.60131
## 2   2013    2 0.2536111     317   quality      6 23.65827
## 3   2013    2 0.2536111     317 preferred      0  0.00000
## 208 2016   19 0.2322222     317 preferred      0  0.00000
## 209 2016   19 0.2322222     317 memorable      0  0.00000
## 210 2016   19 0.2322222     317    trophy      0  0.00000
```

```
str(cpe)
```

```
## 'data.frame':    210 obs. of  7 variables:
##  $ Year   : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2013 ...
##  $ Site   : int  2 2 2 2 2 4 4 4 4 4 ...
##  $ effort : num  0.254 0.254 0.254 0.254 0.254 ...
##  $ Species: int  317 317 317 317 317 317 317 317 317 317 ...
##  $ gcat   : Factor w/ 5 levels "memorable","preferred",..: 4 3 2 1 5 4 3 2 1 5 ...
##  $ caught : int  7 6 0 0 0 9 7 0 0 0 ...
##  $ cpe.hr : num  27.6 23.7 0 0 0 ...
```

```
cpe2 <- aggregate(cpe.hr ~ Year + gcat, data = cpe, FUN = mean)

cpe2
```

```
##    Year      gcat    cpe.hr
## 1  2013 memorable  0.000000
## 2  2014 memorable  0.000000
## 3  2015 memorable  0.000000
## 4  2016 memorable  0.000000
## 5  2013 preferred  5.986703
## 6  2014 preferred  4.938368
## 7  2015 preferred  6.201574
## 8  2016 preferred  2.419355
## 9  2013   quality 17.259044
## 10 2014   quality 16.879829
## 11 2015   quality 15.024400
## 12 2016   quality 14.509571
## 13 2013     stock 17.527035
## 14 2014     stock 19.249018
## 15 2015     stock  5.650490
## 16 2016     stock 17.013989
```

```
## 17 2013    trophy  0.000000
## 18 2014    trophy  0.000000
## 19 2015    trophy  0.000000
## 20 2016    trophy  0.000000
```

```
lmbs %<>% mutate(gcatQ=mapvalues(gcat,
                                 from=c("stock","quality","preferred"),
                                 to=c("quality-", "quality+", "quality+")),
                 gcatQ=droplevels(gcatQ))

(lmb.LFQ <- xtabs(~Year+gcatQ,data = lmbs))
```

```
##       gcatQ
## Year   quality+ quality-
##   2013       56       41
##   2014       75       65
##   2015       53       14
##   2016       54       52
```

```
chisq.test(lmb.LFQ)
```

```
##
##  Pearson's Chi-squared test
##
## data:  lmb.LFQ
## X-squared = 15.552, df = 3, p-value = 0.001401
```

Fix if anyone wants to see 2017

Compare PSD-Q between years 2013 - 2017

```
lmbs %<>% mutate(gcatQ=mapvalues(gcat,
                                 from=c("stock","quality","preferred"),
                                 to=c("quality-", "quality+", "quality+")),
                 gcatQ=droplevels(gcatQ))

(lmb.LFQ <- xtabs(~Year+gcatQ,data = lmbs))

chisq.test(lmb.LFQ)
```

```
(ps.Q <- c(chisq.test(lmb.LFQ[c(1,2),])$p.value,  ### 2013-2014
           chisq.test(lmb.LFQ[c(1,3),])$p.value,  ### 2013-2015
           chisq.test(lmb.LFQ[c(1,4),])$p.value,  ### 2013-2016
           chisq.test(lmb.LFQ[c(1,5),])$p.value,  ### 2013-2017
           chisq.test(lmb.LFQ[c(2,3),])$p.value,  ### 2014-2015
           chisq.test(lmb.LFQ[c(2,4),])$p.value,  ### 2014-2016
           chisq.test(lmb.LFQ[c(2,5),])$p.value,  ### 2014-2017
           chisq.test(lmb.LFQ[c(3,4),])$p.value,  ### 2015-2016
           chisq.test(lmb.LFQ[c(3,5),])$p.value,  ### 2015-2017
           chisq.test(lmb.LFQ[c(4,5),])$p.value)) ### 2016-2017

(p.val.Q <- p.adjust(ps.Q))
```

The PSD-Q of largemouth bass is different for at least one of the years during 2013 - 2017 (Chi-Squared, $X^2 = 16.815$, df = 4, p = 0.0021). The adjusted p-values show a *significant difference* in PSD-Q between years 2014 - 2015 (p = 0.0064) and 2015 - 2016 (p = 0.0046). The PSD-Q is not different between any other years. However 2013 and 2015 may be different (p = 0.0675).

Compare PSD-P between years 2013 - 2017

```
lmbs %<>% mutate(gcatP=mapvalues(gcat,
                                  from=c("stock","quality","preferred"),
                                  to=c("preferred-", "preferred-", "preferred+")),
                 gcatP=droplevels(gcatP))

(lmb.LFP <- xtabs(~Year+gcatP,data = lmbs))

chisq.test(lmb.LFP)
```

```
(ps.P <- c(chisq.test(lmb.LFP[c(1,2),])$p.value,  ### 2013-2014
           chisq.test(lmb.LFP[c(1,3),])$p.value,  ### 2013-2015
           chisq.test(lmb.LFP[c(1,4),])$p.value,  ### 2013-2016
           chisq.test(lmb.LFP[c(1,5),])$p.value,  ### 2013-2017
           chisq.test(lmb.LFP[c(2,3),])$p.value,  ### 2014-2015
           chisq.test(lmb.LFP[c(2,4),])$p.value,  ### 2014-2016
           chisq.test(lmb.LFP[c(2,5),])$p.value,  ### 2014-2017
           chisq.test(lmb.LFP[c(3,4),])$p.value,  ### 2015-2016
           chisq.test(lmb.LFP[c(3,5),])$p.value,  ### 2015-2017
           chisq.test(lmb.LFP[c(4,5),])$p.value)) ### 2016-2017

(p.val.P <- p.adjust(ps.P))
```

The PSD-P of largemouth bass is not different for any years during 2013 - 2017 (Chi-Squared, $X^2 = 8.26$, df = 4, p = 0.08). The adjusted p-values show no difference in the PSD-P between years (2013 - 2017). T